The Computational Challenges of Medical Imaging

Study Leader:
Christopher Stubbs

Contributors:
Michael Brenner
Alvin Despain
Robert Henderson
Darrell Long
William Press
John Tonry
Peter Weinberger

February 2004

JSR-03-300

JASON
The MITRE Corporation
7515 Colshire Drive
McLean, Virginia 22102-7508
(703) 883-6997

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	Februarv2004	

4. TITLE AND SUBTITLE

The Computational Challenges of Medical Imaging

5. FUNDING NUMBERS

13039021-DC

6. AUTHOR(S)

Christopher Stubbs, et al.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

The MITRE Corporation
JASON Program Office – T130
7515 Colshire Drive
McLean, Virginia 22102

8. PERFORMING ORGANIZATION REPORT NUMBER

JSR-03-300

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Department of Energy
Office of Science
Washington, DC 20585

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

JSR-03-300

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution unlimited.

12b. DISTRIBUTION CODE

Distribution A

13. ABSTRACT *(Maximum 200 words)*

JASON will undertake a study for the DOE and the NIH National Institute for Bio-medical Imaging and Bio-engineering on the role of computation (broadly defined to include raw computational capabilities, mass storage needs, and connectivity) for medical imaging. This study will address the computational requirements in three general areas:

- The fusion of image data of varying modalities, over differing spatial and temporal scales and resolutions.
- The extraction and display of quantitative information, with associated uncertainties.
- Data archiving: raw vs. extracted parameters, metadata standards.

14. SUBJECT TERMS

15. NUMBER OF PAGES

53

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	

Contents

EXECUTIVE SUMMARY

Recent progress in developing and refining the techniques and tools for medical imaging, coupled with the ever-increasing power and cost-effectiveness of computational platforms and mass storage, has led to tremendous progress in both research and clinical biomedical imaging applications. JASON was asked to consider what computational needs were likely to arise (with a focus on the next 5 years), and to suggest an effective strategy for addressing these needs. The study's task statement is reproduced below.

Computation for Medical Image Processing: Task Statement

JASON will undertake a study for the DOE and the NIH National Institute for Bio-medical Imaging and Bio-engineering on the role of computation (broadly defined to include raw computational capabilities, mass storage needs, and connectivity) for medical imaging. This study will address the computational requirements in three general areas:

- *The fusion of image data of varying modalities, over differing spatial and temporal scales and resolutions.*
- *The extraction and display of quantitative information, with associated uncertainties.*
- *Data archiving: raw vs. extracted parameters, metadata standards.*

JASON will assess the present status of computational, storage and connectivity needs for existing tools and techniques, and will project likely computational demands for the future. The imaging systems under consideration include both diagnostic and real-time clinical tools.

We are cognizant of other recent study reports that pertain to biomedical computing, notably the Biomedical Information Science and Technology

Initiative (BISTI) report (1), the Coalition for Advanced Scientific Computing (CASC) report (2), and the President's Information Technology Advisory Committee (PITAC) report (3). Our focus was considerably narrower than those adopted in these prior studies; we have concentrated on the anticipated computational needs that are specific to medical imaging.

Although the bulk of this report deals with a five-year outlook, we have included some thoughts on possible long-term opportunities that would arise from applying Petaflop scale computing in the medical imaging arena.

Findings

Our study team was impressed with the sophistication and the approaches being pursued by the medical imaging community. Challenges such as geometrical registration of images of differing modality, for example "lining up" a CAT scan with an MRI image, are being undertaken with a powerful blend of applied mathematics and computational resources.

The current practice in typical clinical applications of biomedical imaging is to present 2-d image data (after suitable preprocessing) to a human who carries out a qualitative assessment based on expert judgment. After being interpreted by an expert physician, the images are archived as part of the hospital's patient records system.

With contemporary computing capabilities, near real-time processing of clinical biomedical images into a form suitable for qualitative analysis is commonplace. The data volumes to be archived do not present a major challenge to large capacity mass storage systems.

Turning now to the biomedical imaging research community, we did not encounter many instances of image analysis problems that were facing major computational throughput bottlenecks. The volumes of images being acquired do not overflow contemporary data storage resources. Multi-CPU parallel computer clusters and Terabyte-scale disks arrays now carry price tags in the tens of thousands of dollars, and as long as adequate financial resources are provided to this community, the hardware should be able to

keep up with image processing pipelines that produce images appropriate for qualitative analysis.

So why can't today's physicians, after acquiring diagnostic images of a patient, query a database to extract similar past cases from a database, with quantitative image properties and fused image-plus-analysis data from differing imaging modalities, including an on-the-fly determination of the historical effectiveness of different treatments for such cases?

Some of the barriers to implementing this vision include establishing metrics for gauging similarities and differences in complex biological images, having a community-wide set of metadata standards for both images and database structures, and incorporating the quantitative analysis of biomedical images into the culture of clinicians.

Our study team has attempted to identify a number of steps that the DOE and NIH could take to address what we see as the major outstanding impediments to progress, and these are summarized in the next section. The main body of the report provides further detail on our findings and suggestions, and responds to the sponsor's request for our suggestions for areas where additional investment might be the most effective.

Recommendations

1. Implement the BISTI report recommendations. In particular their recommendation number 4, pertaining to the availability of a hierarchy of computing platforms for the biological community, is essential to continued progress in biomedical imaging. Moore's Law only benefits those who continue to invest in computing hardware!

2. Calibrate. The lack of a concrete geometrical registration hampers image fusion, and uncalibrated absorbtion or other information hamper quantitative interpretation of biomedical images. We encourage working towards distribution of 3-d standards for geometrical registration frames, incorporating calibration as in integral part of each measure-

ment, and appending the calibration information to all raw data files.

3. Cultivate an open-access and open-source approach to biomedical imaging data sets and analysis algorithms. There are significant cultural impediments within the biomedical imaging community to the sharing of images and algorithms. Furthermore, there are no common 'test problems' against which new algorithms can be tested. We advocate addressing these issues by nurturing the sharing of both code and data. One specific possibility is given in the following recommendation.

4. Establish an open ("BioLena") data set, which all researchers can use to test algorithms and techniques. Implementing prototype metadata standards, NIBIB could act as curators, allowing apples-to-apples comparisons and industry standard test problems.

5. Promote computer-assisted qualitative analysis of biomedical images in the clinical arena. This intermediate step strikes us an achievable near-term goal along the path towards eventual automated quantitative analysis of biomedical images.

6. Develop appropriate database technology, and select and evaluate demonstration projects. We see the database challenges associated with biomedical image exploitation as a major technical bottleneck in the coming years, but one which can be somewhat averted if appropriate steps are taken now.

7. Establish a succession of "Grand Challenge Problems in Biomedical Imaging" to stimulate technical progress on the roadblock issues listed above. This can also galvanize collaborations between the biomedical community, mathematicians and computational and database scientists. Example problems include:

    - Map-the-Phantom – Construct a full-scale anatomical model (thoracic, cerebral?) and invite teams to acquire images and then provide their best quantitative, distortion-corrected reconstruction of

the interior structure of the model. Kudos to those who produce the highest fidelity data set.

- Multi-scale integration – Functional imaging of a biological process, from molecular to physiology. Examples are the cardio and brain efforts already under way.
- Time-to-solution challenge – Pick an imaging methodology and problem. Points for whoever can port their analysis toolkit to a standard platform and get an acceptable answer the fastest. Also points for the "best" answer.
- Quantitative Change Detection Challenge – Given a temporal sequence of images, some with actual clinical data and others with features inserted "by hand", identify and quantify the evolution of the changes.
- Joint analysis challenge – Use raw data from multiple modalities to improve the fidelity of image generation.
- Quantitative Diagnosis challenge – Use parametric descriptions of image features of interest to achieve detection, diagnosis or differential diagnosis, as appropriate.
- Multimode image integration challenge – Produce the best registered set of images, with common data structure and access tools, from images obtained with diverse methods.
- Best Merged Image-plus-catalog data structure, with query tools and comparison metrics for images.

8. Begin the process of considering the potential of using what we presently consider super-computing in the biomedical imaging arena. Today's supercomputer is tomorrow's desktop machine, and this may open up totally new approaches to the interpretation of biomedical images.

1 INTRODUCTION

Computing in the biosciences is a major endeavor, encompassing everything from protein folding to genetic databases to the information technology associated with patient medical records. Our task was to look into the likely computational requirements for a narrow slice of biomedical computing, namely the CPU, storage, software and connectivity requirements needed to digest, exploit and archive the images produced by the clinical and research biomedical imaging communities. Even this subset of biomedical computing covers a wide span of activity. This report makes broad observations and recommendations, based on our study team's findings and experience. While there are undoubtedly specific counterexamples to many of the points we make, we contend that some general trends do emerge and that there are specific opportunities for high-impact investment by the NIH and the DOE.

For the purposes of this study we define biomedical imaging as the collection of methods used to produce 2 or 3 dimensional representations of physical properties of systems of biological interest. This includes optical and electron microscopy, Xray (CT) and electron tomographic (ET) imaging, Positron Emission Tomography (PET), Single Photon Emission Computerized Tomography (SPECT), ultrasound, Nuclear Magnetic Resonance (MRI) and Electrographic and Magentographic Encephalography (EEG and MEG).

These techniques have different domains of applicability, ranging from probing biological systems at molecular scales to full-body scans. Although there are exceptions, presently most clinical applications utilize organ to limb scale images, while biomedical imaging at smaller scales currently mostly supports basic science research. Figure 1 presents a map of biomedical imaging in the context of applications and characteristic length scales. As discussed below, (1) increasing the clinical applications of imaging at smaller scales, and (2) spanning many decades of length scales to improve our understanding of biological processes do involve computational challenges.

It is important to recognize that there are two distinct biomedical imag-

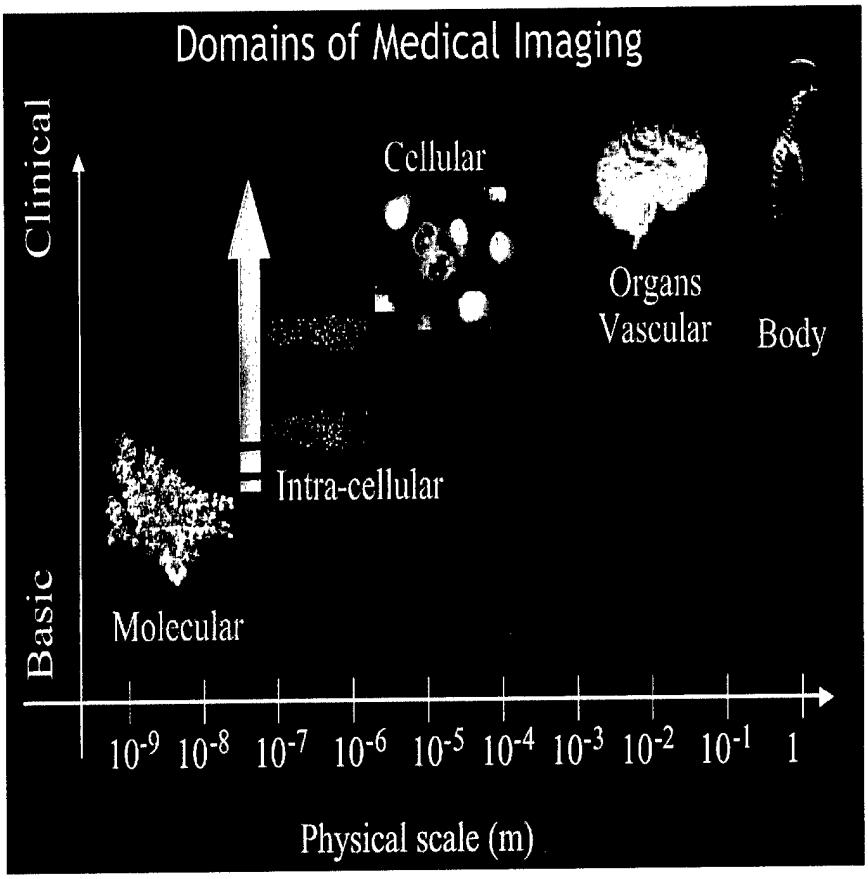

Figure 1: Medical Imaging domains of applicability. The horizontal axis is a logarithmic representation of characteristic length scales, while the vertical axis reflects the distribution of applications, from basic research to clinical applications. A major goal should be to shift current research approaches "upwards" on the plot, towards routine clinical use, when appropriate.

ing communities, with some overlap between them: 1) the clinicians, who are most interested in maximizing accurate information of medical interest at the lowest possible cost, typically using commercial hardware and software, and 2) the biomedical imaging research community, who frequently have a closer relationship to the acquisition hardware and analysis software, and who can often tolerate longer latencies in image processing times.

Overall, the JASON study team was impressed with the existing ongoing efforts in the community's approach to the computational challenges of biomedical imaging. We saw a nice combination of applied mathemat-

ics, medical physics, and innovative computational techniques that were well coupled to the biological problems at hand. We did not hear the biomedical imaging community express frustrations at lack of adequate computational throughput, or even for lack of disk space. For research applications, computing at the scale of a high-end cluster of Linux workstations seems well suited to handle current image processing needs. There is of course a worry that this represents a 'selection effect', in that the algorithms currently being used are precisely those that can return an answer in a tolerable amount of wall clock time, but we did not sense that there are significant unexploited opportunities that remain dormant for lack of adequate computational throughput. With the cost of both CPU power and disk storage falling rapidly, we do not expect that either of these will produce bottlenecks in the 5 years ahead. (We do note later in this document the potential benefits of applying supercomputer technology to medical imaging analysis, however.)

So if the hardware is not a major source of concern, what *is* hard about computing in the biomedical imaging arena? We highlight the following list of issues that are (or are likely to become) impediments to capitalizing on the ongoing technical developments in both imaging techniques and computing capabilities:

- The evolution from qualitative to quantitative interpretation of biomedical images is hampered by the fact that the questions are ill-posed. Computing the morphology of shapes is a tough problem. This difficulty is accentuated by the fact that the heritage for generating well calibrated image data sets is not particularly strong.

- The lack of a common set of metrics, and the absence of a standard set of test images/cases, makes it difficult to quantitatively compare different techniques and algorithms.

- As the imaging techniques used by the research community at the subcellular and molecular level make a transition into clinical applications, the challenge of fusing information across length scales, phenomenology, and imaging modalities must be confronted.

- A major obstacle that we foresee is the inability of current database technologies to easily accommodate images as intrinsic database objects. Most high-volume image archives that are presently linked to SQL-compatible databases do not contain the images themselves as data entities, but rather the databases typically contain links to where the images are stored on disk. As outlined later in this report, we see this as an area ripe for investment.

- A related issue pertains to the lack of metadata standards. This will soon become a significant impediment to interoperability across data structures, and to the effective sharing of data between subdisciplines in the biomedical imaging community.

- There are significant cultural obstacles pertaining to the sharing of biomedical image data and algorithm source code, which have in our view hampered progress in this discipline. We comment on these issues, and potential ways to start overcoming them, in the sections that follow.

We received briefings from a diverse cross-section of the medical imaging community. The speakers and their institutional affiliations are listed in Table 1. In addition, we benefited from extensive conversations with other members of the biomedical imaging profession. We are very grateful to all of these individuals for taking the time to share their viewpoints, concerns and suggestions with us.

The structure of this report largely traces the flow of information in a medical imaging application. We start in Section 2 with the initial step of converting raw data into images, i.e. going from bits into pixels. The resulting images are now typically presented to experts (physicians) who use their extensive experience and professional judgment to extract knowledge from the pictures. Their interpretation is usually presented as a narrative qualitative appraisal, sometimes even only comprising a single bit of information (yes/no). We note for future reference that this analysis constitutes a very significant reduction in data volume, from a digital image to a few succinct bits of pertinent extracted information.

Table 1: Study Briefers

Speaker	Affiliation
Richard Leahy, Ph.D.	NeuroImaging Research Group University of Southern California Los Angeles CA
Chris Johnson, Ph.D.	Director, Scientific Computing and Imaging Institute University of Utah Salt Lake City, Utah
Michael Miller, Ph.D.	Director, Center for Imaging Science The Johns Hopkins University Baltimore MD
Mark Ellisman	Director, Center for Imaging and Microscopy Research University of California, San Diego
Larry Frank	Center for Functional MRI Imaging University of California, San Diego
Michael Vannier	Chair, Department of Radiology University of Iowa
Richard Martino	Director, Division of Computational Bioscience NIH Center for Information Technology, Bethesda MD
Judith Niland	Division of Information Sciences City of Hope Hospital, Pasadena CA

We will explore what is needed in order to shift from this qualitative interpretation to a fully quantitative analysis of medical images. This includes the transition from pictures to numbers, using parametric analysis techniques, discussed in Section 3, and the extraction of knowledge and understanding from these image parameters. This is discussed in Sections 4 and 5. Issues relating to transfer of information and connectivity are discussed in Section 6. We have significant concerns that pertain to data and code access, which are discussed in Section 7. Considerations of the application of truly high end computing to medical image analysis is presented in Section 8. We close with our recommendations and conclusions in Section 9.

2 THE ANALYSIS OF RAW DATA: FROM BITS TO PICTURES

A common feature of all biomedical imaging techniques involves the transformation of raw digital data into a image. Solving the inverse problem for CAT scans is one example of the type of analysis required. Another example is the forward modeling used for MEG analysis. The refinement of the tools and techniques used to produce images of high fidelity is an ongoing field of research, and one that certainly merits continued support. We note that in many fields the development of more efficient and effective algorithms accounts for as much increased analysis throughput as hardware improvements.

The analyses needed to convert from raw data into images often come to us as ill-posed problems, with incomplete, grainy or noisy data. A major challenge in the coming years will be to take the experimental uncertainties, and the *assumptions* made in the analysis, and propagate them forward through the visualization and interpretation stages.

2.1 Raw Data Volume and Data Rates – Not A Major Limitation

Using a nominal rate of 32GB/hour, a single MRI machine generates approximately 11TB/year of raw data. Certainly handling 11TB of data is not difficult technically; these data can be readily stored in a few RAID arrays of disks, in a rack with high throughput interconnects. Providing local (same-building) access to comprehensive image archives is also not a major technical challenge. We do note that patient record confidentiality and access permission issues may be a hurdle, but one that will surely be surmounted in the coming few years.

Given adequate financial resources to acquire the requisite hardware, and sufficient system administration support for the platforms and disk

farms, major municipal hospitals should be able to generate and maintain image archives for their patients.

As discussed below in Section 6, however, transferring large image archives around on the net will likely be a significant bottleneck in the decade ahead.

2.2 Converting from Raw Data to Images: Inversion Techniques and Forward Modeling

One class of medical image generation uses inverse techniques to generate an image from the raw data. A classic example is the inversion used to extract a model of the body using x-ray transmission measurements as a function of angle. Although much progress has been made on devising clever techniques for CAT scan analysis, it remains the case that the resulting images do *not* typically contain uncertainties associated with the analysis technique, the assumptions, or the even the signal-to-noise ratio of the raw data.

If the imaging method were perfect, producing images would be straightforward and error free. For example, tomographic reconstruction requires inverting a Radon transform. The mathematical properties of this transform have long been understood.

In practice, difficulties arise because of uncertainties in the imaging method. Uncertainties are caused by a myriad of factors, including (a) systematic errors inherent in the imaging method (e.g. the local magnetic field distortions in MRI typically introduce significant, patient dependent, uncertainties); (b) image contrast limitations (e.g. magnetic resonance does not provide good contrast for bone, while CT scans do not provide good contrast for soft tissues); (c) movement of the patient during the scans; (d) aliasing effects and (e)multiple scattering effects. Each of these factors introduces uncertainty into the imaging data, which in turn causes the inversion of the imaging transform to be mathematically and computationally ill-posed.

The only way to invert an incomplete, error laden, imaging transform

is to make use of some statistical model for the missing and uncertain degrees of freedom. Depending on the level of detail in the statistical model, this can be a computationally demanding task. For example, in particle physics/astrophysics running the calibration statistical models takes up the bulk of the computing time. This approach is not common practice in the medical imaging community.

Recent approaches have lowered the dimensionality of the reconstruction problem by positing that the imaged object is composed of distinct materials with known material properties. Under this assumption, the inversion problem need only solve for the interface between the distinct regions. Although progress is clearly being made in this direction, attention must be paid to what the errors are, even under the assumption of perfect interface inversion. For example, it is well known that tissues are not isotropic materials, and the anisotropies must affect the data obtained.

The other major class of imaging challenges that arises in medical imaging is the problem of *forward modeling*; using a set of measurements that even under ideal circumstances give incomplete information to exactly reconstruct the image. The major exemplars of this class of problems are magneto encephalography (MEG) and electroencephalography (EEG). These techniques require measuring the electrical potential or magnetic field at a finite set of sensor locations distributed over a region of the body. From these data, the task is to reconstruct the charge/current distribution inside the region imaged. Even in a perfectly characterized material this problem is ill posed: determining (for example) the charge distribution inside a body requires knowing the electric potential on the entire surface of the body. The charge distribution that reproduces a finite number of measurements of the potential on the surface of a body is not unique. Said differently, in a perfectly characterized material, there is a set of charge distributions that is quantitatively consistent with a given set of measurements. This dispersion in the set of charge distributions gives the error in the interpretation of the data.

Medical imaging brings several additional complications: First, the material properties (dielectric and conductivities) are generally uncharacterized.

Second, for clinical imaging, the shape of the surface over which the data is taken is not characterized. Even if the surface potential were measured exactly everywhere, each of these factors would be an uncertainty in the deduced interior charge/current distribution. For the errors associated with either of these issues to be reduced, it is clear that EEG/MEG need to be combined with other imaging modalities that are capable of measuring the shapes of surfaces and the shapes of regions with different material properties. Such work is the focus of current research activity, though the field has a long way to go to fully understand and accommodate the uncertainties from both the data and the image-generation techniques.

Ongoing support of algorithmic development is well warranted, but would be enhanced by increasing access to both test images and algorithms, as discussed in Section 7.

2.3 Reducing the Uncertainties in Image Generation by Simultaneous Joint Analysis

A major hurdle that must be surmounted is to reduce the uncertainties described above. The community is carrying out various research directions in this regard, mainly with the view towards combining complementary techniques together. For example, combining MRI with CT scan would allow simultaneous visualization of soft tissues and bones, if the two types of images could be calibrated against each other. Combining MEG with MRI gives information about the precise surface of the head to be combined with inversion calculations for the MEG.

We note that a joint analysis of the *raw data* is fundamentally different from the challenge of fusing information from images, at a post-analysis stage.

Although it is clear that the combining of different methodologies will lead to more and better information, for quantitative metrics to be developed it is imperative that the data from each of the imaging methods be calibrated

to a common reference standard so that they can be used together without additional calibration error.

2.4 The Merits of Calibration

Current efforts on computer analysis of medical imaging are focused on trying to correlate morphology with function of disease. The morphology is generally defined by surfaces which can be distinguished by discontinuities in density or other properties. Thus, the real information being extracted lies in spatially localized differences, and depends on the smoothness of the performance of the imaging systems.

While we are confident that this will yield results, we are not certain on what scale the results will emerge. It is possible that gross anatomical differences in size and shape in parts of organs will not correlate well with function, but that these correlations will not emerge until the cellular or even molecular level.

We were briefed on studies of the hippocampus which tried to correlate shape with schizophrenia. Apparently efforts in that direction were significantly hampered because different MRI machines have spatial distortions which are on the same scale as the shape changes that the investigators were measuring. This made it impossible to fuse data from different machines, and moreover we were told that routine machine maintenance will cause enough distortion change to hamper research along these lines.

We applaud the wonderful accuracy that manufacturers have achieved with their MRI and CAT machines, but at the same time we feel that some of this effort is misplaced. A chest xray will be directly examined by a radiologist without computer enhancement, and therefore must convey the information that the diagnostician needs directly.

However, the raw data from MRI or CAT scans are not useful without computer processing, and so the most important attribute that a machine should have is stability. If a machine is stable, it can be used to scan a

standard target (a cube, for example, with struts of known location and density), and then these reference data can be used as part of the normal computer processing to produce an image which has no distortion at all. We imagine that the first and last use each day of an MRI or CAT machine might be a scan of the standard cube.

Such routine calibration data would convey a technical advantage, of course, in that it monitors the health and accuracy of the machine, and the computer analysis of it could determine whether the machine is operating within specification, or whether it should be serviced.

The most important aspects of incorporating such calibration data into the analysis of biological scans are that

- It permits easy and accurate fusion of many different data sets
- It provides an absolute calibration of size, density, etc.
- It opens the possibility for extremely sensitive temporal studies of the same subject.

This last aspect is one which deserves attention. If two MRI scans of a given subject, possibly taken months apart, were accurately calibrated in terms of position and density, they could be registered and subtracted. The registration displacement field would be an accurate measure of any spatial changes which occurred during the intervening time (swelling around a tumor, for example), and the density changes which would be revealed with great sensitivity might also have significant diagnostic value.

2.5 Enhanced Visualization of Biomedical Images – Computer-Assisted Qualitative Analysis

At present the generation and display of clinical medical images is tailored to support qualitative analysis by physicians. The radiologists draw upon their training and personal experience to interpret imaging data, and

to arrive at a diagnosis, or differential diagnosis. This often involves comparing an image with the clinician's *recollection* of what normal or pathological features look like. In many cases, a patient's imagery is compared not with earlier images but with written reports about prior images. Even when an image history is available, the comparisons are not quantitative but are simple side-by-side comparisons made by the radiologist. In the case of comparing images with written reports, it is important to note that a different radiologist may be making the comparison, and that due to its subjective nature, important features may have been missed. A small feature discounted by the first radiologist may now be manifested as disease in the patient. Finally, the images are considered in the context of other case-specific information: patient age, signs and symptoms, medical history, etc.

2.6 A Valuable Near Term Opportunity

It strikes us that the present clinical procedures could be enhanced, using computers and archived images to improve the performance of qualitative biomedical image analysis. There are a variety of ways this could be done, two of which we list below:

- Using archived images and computer-assisted access to provide relevant comparison images and information. We envision a system in which the physician is presented with a montage of relevant comparison images, reflecting stages of disease progression, and (when appropriate) examples of benign physiological anomalies that are commonly mistaken as disease. The evaluation could then be carried out with real-time access to relevant comparison images. The determination of an *appropriate* comparison set of images is a challenge, but not an insurmountable one. We will return to this issue in the section on databases.

- Having an analysis program draw the physician's attention to image features of potential interest. There is of course the potential problem of having radiologists become overly reliant upon this, and potentially

missing important features. This could be avoided by having the computer analysis take place after the radiologist has made an appraisal, as a backstop to the human interpretation.

This middle ground, using computational resources to enhance qualitative image analysis, strikes us as an effective way to build technical and cultural bridges into the era of full quantitative analysis that is surely in our future.

2.7 The Representation of Uncertainties in Medical Images

The computer graphics and scientific visualization communities have conducted extensive research in the visualization of uncertainties. That work should be leveraged instead of reinventing it. Similarly, the computer graphics and scientific visualization community have significant expertise in dealing with isosurfaces and textures. It seems that this body of research could be leveraged by the medical imaging community. Research collaboration with computer scientists in these fields should be expanded significantly.

Providing visualization of uncertainties is not the hard part here– rather obtaining and propagating the underlying uncertainties is the real challenge. Presently, medical images do not contain uncertainty information. These uncertainties can arise from fundamental limitations in the measurements (Poisson noise, etc.) or from uncertainties due to the image-generation technique. Clearly, the community must first decide that uncertainties are an integral part of medical images. Then, different techniques for visualization and representation can be evaluated.

Incorporating uncertainties as an integral part of biomedical images is of course a necessary prerequisite to being able to carry out fully quantitative analyses of these data.

3 THE CHALLENGES OF QUANTITATIVE IMAGE ANALYSIS: EXTRACTING NUMBERS FROM PICTURES

As outlined in the previous section, people are accustomed to looking at pictures, and other 2 dimensional representations of information. In clinical applications the primary product of medical imaging is a 2-d image. There is a long tradition in radiology of deriving very useful clinical information from the qualitative examination of such images. This qualitative approach, with expert judgment by physicians leading to narrative descriptions of findings, does not do justice to the rich information contained in images obtained from contemporary imaging systems, and limits the physician's ability to quantitatively express the clinically essential information in a succinct way.

3.1 The Merits of Quantitative Analysis

There are numerous benefits to be reaped from moving towards a more quantitative exploitation of medical images: Monitoring the progress of a medical condition, and ascertaining its response to therapies, would be enhanced if the community had reliable and effective quantitative tools. Comparisons with archived images of comparable cases would be facilitated with quantitative tools, and quantitative descriptors are likely to play in key role in identifying relevant image data.

There are some examples of quantitative analysis of medical images in a clinical setting, such as physiological measurements made on ultrasound images, but this is presently the exception rather than the rule. The quantitative analysis of ultrasound images is an encouraging case where a new technique was rapidly adopted by the clinical community once its value was clearly demonstrated.

One approach to quantitative analysis of biomedical images involves extracting from (2-d or 3-d images) a parametric description of the morphology

of objects of interest in the frames. This would involve, for example, automated recognition of physiological features in images (tibia, femur...), and a means for summarizing their properties with a handful of numbers (length, width, density...).

3.2 Change Analysis with Image Subtraction

An alternative approach would be to carry out a *differential* analysis of a succession of images. We wondered whether the image subtraction schemes used (7) in astronomy to detect change might have application in this arena as well. We will not pursue this further here, but we do advocate evaluating this approach. Figure 2 shows the power of literally subtracting images in order to highlight changes. Software currently used in the astronomical com-

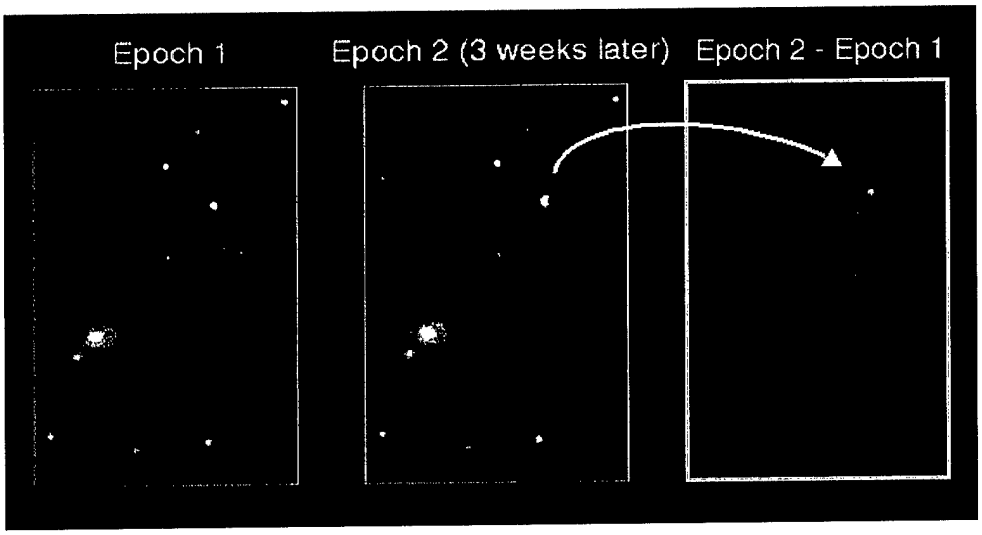

Figure 2: An example of image difference analysis. The figure shows two images taken at different times in the left and center panels. The right hand panel shows the pixel-by-pixel difference in the images, in this case highlighting a supernova. This approach may prove fruitful in the analysis of biomedical images as well. (Image courtesy of the High-z Supernova Team.)

munity can compensate for geometrical distortions, as well as additive and multiplicative scaling between the 2 images before carrying out the subtractions.

Monitoring the progress of a medical condition, and ascertaining its response to therapies, would be enhanced if the community had reliable and effective quantitative tools for evaluating medical images. Such tools would allow the tracking of precise anatomical changes within a given patient over time, and also within patient populations. Such quantitative tools would invite the development of imaging metrics, to track conditions and manage risks. Current quantitative metrics for assessing risks exist throughout medicine, though are notably lacking in many modern imaging technologies.

Present research is motivated by the desire to map out boundaries, surfaces and volumes in biomedical images. This approach follows the presently prevalent notion that pathology is manifested in gross anatomical abnormalities, and we note that this will likely evolve to include more subtle chemical, physical and biomolecular evidence of disease and injury. As this understanding progresses, we can look forward to biomedical imaging modalities that will provide quantitative diagnostic information.

3.3 Why Is Quantitative Image Analysis So Difficult?

With all the effort expended on biomedical imaging technology and analysis, why is the quantitative analysis of this imagery not commonplace? We consider there to be a number of reasons why we have not progressed to its obvious conclusion. Calculational capability is not the limiting factor. Rather, the difficulties include the fact that the extraction of the features of interest is an intrinsically ill-posed problem. While physicians, using a vision system that has been honed by many generations of human evolution, can sift the uninteresting from the informative, it is *very* hard to teach a computer to do the same.

Developing quantitative descriptors of medical images requires not only finding ways to extract a parametric description of the morphology and texture of objects of interest from 2d or 3d images, but also developing measures of *uncertainties* in the reported description. In many of the imaging modalities currently in use, the uncertainties are substantial, depending on

details of the patient, the imaging device, and the manner in which images are acquired. Without an accurate understanding of these uncertainties and a way of representing them, quantitative analysis of images is impossible. This should motivate not only an effort to increase the reliable calibration of medical images, but also the propagation of uncertainties through the entire image analysis pipeline.

Additionally, the pedigree of the extracted features must be retained. This requires tracking and archiving the raw image data, the code used to generate an image, the code used to extract feature parameters, etc. This would ideally all be stored in a self-describing data structure that is intimately linked to a version-controlled code bank, as illustrated in Figure 3.

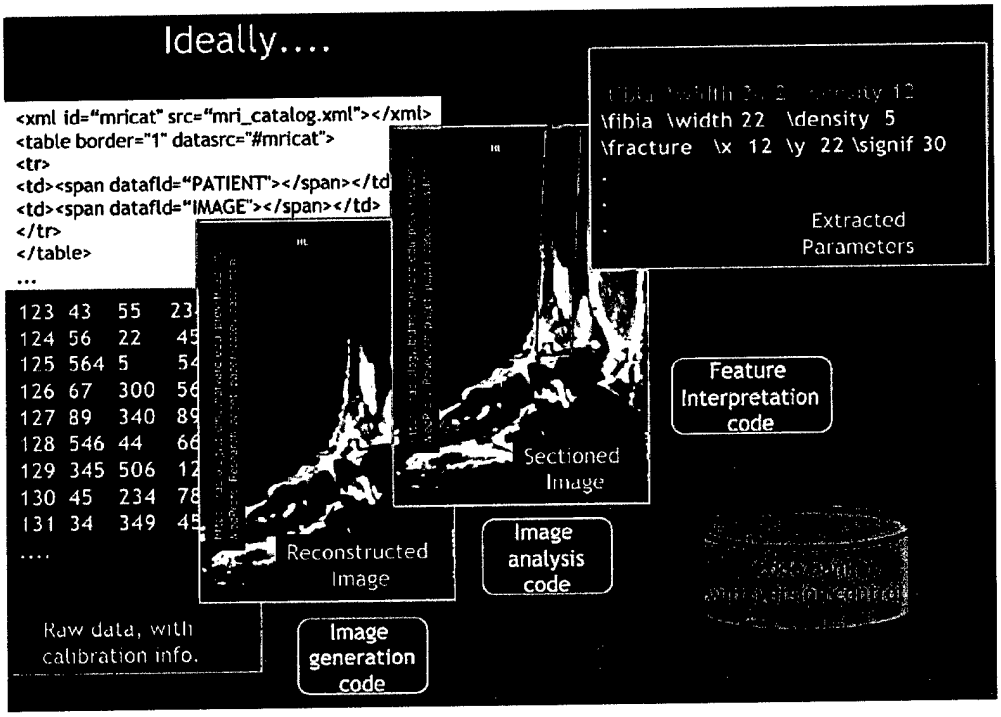

Figure 3: An Aspiration. An integrated image-and-analysis self-describing structure would link raw data, generated images, and an extracted feature catalog with the version-controlled code bank used in the analysis. The full pedigree of the data and code version would be retained in an integrated metadata structure.

Even once a parametric description of image data has been extracted, these numbers must assessed in comparison to other cataloged numbers,

spanning the range from "normal" to "pathological" in order to obtain a clinical appraisal of value. This comparison will necessarily involve a consideration of the specific case history, which is probably best represented as a set of Baysian priors. This challenge is considered in the next section.

4 INTERPRETATION: FROM NUMBERS TO KNOWLEDGE

Once a parametric description of image features of interest has been obtained, the goal is to turn this into useful biomedical information and insight. This will require a comparison with either (1) the relevant parameter history of the patient in question (essentially a differential measurement) or (2) a set of comparison data drawn from a relevant comparison group. This process will clearly benefit from building queryable databases, but we will defer considering that aspect until Section 5.

Any comparison of extracted feature parameters will obviously rely upon having calibrated data with well understood and quantified uncertainties, as any similarities or differences must be considered in the context of their statistical significance. We do not consider the current state-of-the-art in most modalities of biomedical imaging, or in general the analysis of these images, to be at a stage that will support this kind of approach.

4.1 Defining Relevant Comparison Images

The definition of a relevant comparison group is presently done implicitly when a physician interprets a clinical image. The doctor is bringing strong prior probabilities to bear on the problem, based on the patient's clinical history, symptoms, the results of laboratory tests, and other pertinent information. This "data fusion" is a major component of the training that physicians receive, and also draws upon the doctor's personal experience. Moving from this approach to a diagnosis (or differential diagnosis) that is based upon a parametric description of image features will be very challenging. While it is certainly possible to envision constructing a database query that constrains the parameter comparison to, say the typical size of the livers of 12–15 year old girls that live in the Eastern US, we are a long way from being able to carry this out. The availability and low-latency

accessibility of the archived comparison information is a major part of this challenge.

Additional complications include defining appropriate comparison groups for each patient, contending with benign anatomical anomalies, and moving beyond the consideration of surfaces and volumes as the quantities of interest.

We conclude that not only is the extraction of parameters a difficult problem, but that the clinical interpretation of these parameters, by way of comparisons, is far from trivial. So how might the agencies foster progress in this arena? One potential approach would be to pick a demonstration case where the algorithms needed for parameter extraction do not present a major challenge, and where the scope of the comparison group is well defined. This is a good candidate for a "Grand Challenge" in biomedical imaging.

We envision an eventual progression of physician interaction with images and their features. We imagine moving from today's stage of "show me the picture" to being able to extract a subset of images from an archive with commands like "show me all images that contain skull fractures with lengths between 2.5 and 5.0 centimeters", to interactive processing such as "run this new algorithm on all lung images in the archive, and store and compare the results" to eventual natural-language interactions such as "return all images that contain features like this one". This leads us to the interplay between databases, image archives, bandwidth and latency.

5 DATABASES, DATA RETRIEVAL, IMAGE ARCHIVES AND METADATA: A HIGH-LEVERAGE OPPORTUNITY?

5.1 The Potential Value of Sophisticated Databases in Medical Imaging

There is, in our view, a considerable opportunity in developing more sophisticated database tools in support of biomedical imaging. Whether the images themselves are included as intrinsic database objects, or whether the database simply contains pointers to images that reside in an external file structure is an implementation detail. The goal should be to build a tightly integrated data structure that contains

- data pedigree information: code versions, image construction algorithm parameter files, etc.
- image files,
- uncertainty arrays,
- parametric descriptions of detected image features,
- links to patient record data, including updated information about outcomes and progress.

Eventually this field will develop and maintain such database structures that merge calibrated images with extracted parameters (shapes, volumes, etc.). It seems to us essential (and inevitable) that comprehensive biomedical imaging data, both images and extracted parameters, be widely available after addressing patient confidentiality issues.

This will provide a means to access and exploit the increasing volume of medical imaging, in a way that could provide substantial improvements

in patient care. With records that are readily accessible across the nation, a patient who appears in the Denver ER can have their records accessed by that facility, even if they reside half a continent away. Physicians could interact with the aggregate data in order to compare and contrast the case under consideration with the nation's accumulation of such cases.

User–friendly interfaces to these databases will help overcome the risk of building substantial write-once-read-never (WORN) data sets. We see it as essential to build small-scale prototypes, with query efficiency and ease of access as prime considerations.

In an era when digital data seldom outlast the life cycle of proprietary formats and systems, if medical imaging data are stored in compliance with broad meta–data standards, these data will be sustainable over multiple generations of hardware and software evolution. This will require the development and adoption of metadata standards.

5.2 Metadata Standards

If we consider constructing a national archive of medical images for diagnostic and research purposes, this archive will be very large. Depending on the policy for placing images into the (distributed) archive, it could range from a few TB (terabyte) to a PB (petabyte) or more. Such a large archive will require professional management, and high bandwidth links to the researchers and physicians who use it. We should note that given such a large collection of images, it will be impossible in the foreseeable future to use image processing techniques to search through this archive. Searching will have to be done on metadata, and so techniques and metrics for describing features of the images will have to be developed. It should be possible to derive many of these features from the images automatically, and then store them with the image as metadata.

When we consider metadata, and indeed data formats, standards are very important. The reason that computers interoperate on the Internet is

due entirely to the adoption of standards; the reason a cellular telephone works on more than one network is due to adherence to standards. In order for researchers to make effective use of medical imagery data, these data should be put into standard formats that can be read by all researchers. Manufacturers should be encouraged to adopt these standards (we need to try to make a case that this will be to their advantage).

All of the sensor calibrations, the algorithms used to construct the image, the transformations applied to the image, its segmentation and annotations by medical professionals make up much of the metadata of the image. By choosing a standard format, and carefully maintaining this metadata, it becomes a searchable quantity in the database. A query such as "Find all brains with possible aneurysms near the circle of Willis identified using an MR angio with no contrast agent" suddenly becomes possible.

A good example of the use of metadata is the AFNI system (6). The AFNI system carefully annotates medical image data, including calibrations and its lineage and all transformations that have been applied to it. A further improvement would be to adopt a standard metadata description language such as XML. XML is a widely used standard, and since it is well-understood many parsers for it exist, that would ease adoption. By using a standard metadata description, exchange of data and the ability to both track the lineage and changes to that data, as well as make queries against that data would be significantly enhanced. In recent years, database technology has been developed that works well with XML.

It appears that current practice is to use flat text files, and in some cases files encoded in binary formats. The amount of space saved by binary formats is minimal, and not significant given the growth in storage technology. The use of text files improves portability, but the formats are still proprietary and this makes exchanging data with other researchers (or medical professionals) difficult. It is important that the data produced by medical imaging equipment be self-describing, and again a language such as XML seems to be ideal for this task.

XML, or a language like it, could be used to describe data ranging from the raw data returned by the sensors, to the images that are derived from the sensor data. It could describe all calibration coefficients and other parameters as appropriate for the imaging technology. Once an image is constructed, the algorithms and corrections made could be described using XML. As image processing algorithms are applied, each transformation of the image could be appropriately noted.

In the case where the image is segmented, XML could be used to describe the segmentation of the image. Again, we gain the advantage of being able to describe in the image itself how the segmentation was accomplished. Annotations made by medical professionals, such as the identification of features could be kept with the image in the XML. For example, the identification of an aneurysm, its type and location could be made.

The archival situation in the clinical setting is very poor. Archives, when kept, are usually kept as film and not digital images. Due to the nature of the turn-key systems currently sold, images from an older system may not be compatible with those of the new system. The DICOM standard used in the medical imaging industry seems to have serious compatibility problems, and we wonder why yet another standard was thought to be necessary in the presence of so many digital image standards with proven compatibility.

JASON sees the generation of broadly supported metadata standards and the development of appropriate database testbeds as important steps in moving the medical imaging community towards realizing the full potential of the discipline.

6 CONNECTIVITY: PUSHING A RIVER THROUGH A STRAW

Although Terabyte data volumes can be cost-effectively stored on disk, the bandwidth needed to support the exchange of these data sets is not currently available. This can be readily illustrated by estimating the time required to transfer a typical image: a 1K x 1K image at 16 bits comprises 2 MBytes of data. Moving this across a network with a delivered bandwidth of 100 Kbits/sec would take nearly 3 minutes. Implementing a scheme where large image data sets are routinely transferred across the nation will rapidly saturate the existing network capacity.

There is a fundamental mismatch between the image archive size that can be readily stored locally (tens of Terabytes) and what can be transferred across the network (optimistically, perhaps Gigabytes/day). We do applaud initiatives such as the BIRN project(8) that are stepping up to these challenges, but we feel the network infrastructure is not able to support wholesale exchange of large image data sets.

The agencies should take a hard look at nationwide networking capacity, and anticipate the likely evolution of demand from the medical imaging community.

Local networking infrastructure is also an important issue that needs to be addressed. Deploying sufficient infrastructure locally in a building or group of buildings on a campus is not prohibitively expensive, but it is important that the infrastructure be kept up to date on a regular schedule. Currently, that infrastructure should be 1 gigabit Ethernet for data transfer; but a plan should be in place to move to the next generation as soon as it becomes cost effective. The more difficult issue is connectivity among geographically distributed researchers and clinicians. The so-called "last mile problem" has not be inadequately resolved, and so aside from high cost solutions getting sufficient bandwidth remains expensive. Short-lived initiatives to connect clinics and hospitals are not sufficient.

Two possible approaches to overcoming the connectivity gap are 1) image compression and 2) parameter extraction. We understand that liability issues preclude the use of lossy compression for medical images, but there are modest factors to be gained by using lossless compression algorithms. The other approach is to avoid transferring full images, but rather to transfer extracted feature parameters, which is a much smaller data volume.

We do anticipate that network capacity limitations will likely prevent the full benefits of rapid image exchange from being realized, unless steps are taken to increase network throughput, on both the national and local scales.

7 DATA ACCESS AND RELATED CULTURAL ISSUES

The biomedical imaging community does not have a strong heritage of releasing image sets or code, even upon publication. This stands in stark contrast to the approach taken by the molecular biology community, where publication of research papers is contingent upon gene sequences being deposited in an accessible database. The following excerpt (4) from a recent editorial in *Radiology* paints a grim picture –

> *In radiology, where imaging is central to everything we do, published images are neither indexed separately nor retrievable. To make matters worse, most authors decline to share their original source images, preferring to maintain them in private collections. It is impossible to reconstruct the results of published work, since the original source data (e.g., images) are unavailable.*

In apparent recognition of the importance of clarifying its approach to proprietary data, the NIH has issued (5) a Data Access Policy, an excerpt of which reads

> *...Starting with the October 1, 2003 receipt date, investigators submitting an NIH application seeking $500,000 or more in direct costs in any single year are expected to include a plan for data sharing or state why data sharing is not possible...*

Having set the criteria for what constitutes a project whose data are considered of sufficient value to merit a data release plan, the NIH policy then goes on to instruct (5) reviewers to disregard the strength or credibility of the data release plan in assessing the merit of the proposal:

> *Reviewers will not factor the proposed data-sharing plan into the determination of scientific merit or priority score. Program staff*

will be responsible for overseeing the data sharing policy and for assessing the appropriateness and adequacy of the proposed data-sharing plan.

Several efforts to gather and freely distribute biomedical image archives have been attempted, but most have withered due to lack of enthusiasm by researchers. It seems to us that a change in attitude is necessary. One only needs to observe the benefit of data sharing enjoyed by the genome community to see that science is better served by open access to data than by holding those data confidential. Other branches of science have already embraced this goal, for example, if an astronomer is funded by NASA, then in 18 months all images created under that award enter the public domain. NIH has begun with a much weaker model, requiring researchers to develop a data sharing plan as part of their grant applications. Often funding is not including for data sharing, and data sharing usually consists of ad hoc web pages.

The culture of data-hoarding that appears to permeate much of biomedical imaging research strikes us as outdated. It limits progress in the field, and prevents an honest comparison of tools and techniques. Other scientific disciplines have wrestled with the issue of proprietary data rights, and there is a strong trend towards increasing community access to data sets and analysis tools that have been developed with taxpayer funds. We note the thoughtful narrative from NASA on this topic (9). Certainly in astronomy, much of the improvement in open access is a direct result of funding agency policies. In this context, we found the NIH data access provisions to be somewhat less than ideal, in pushing the field towards more open access to image data.

Another limitation is lack of a standard test set of data, so that different algorithms and approaches can be compared. Most research papers that describe new image generation or analysis algorithms present 'before' and 'after' images for qualitative comparison, but the images themselves (let alone the algorithms!) are seldom made available to the community. This makes

it nearly impossible to make a quantitative comparison of the performance of different approaches and algorithms, as there is no common set of test images.

We see considerable merit to the idea of establishing an open-access data archive, conforming to prototype metadata standards, from which the research community could draw example images. Results of various inversion and analysis algorithms could then be *uploaded* to this site (even along with code, if that cultural barrier can ever be breached). This is a chance to push towards an open source/open data ethic.

We encourage the agencies to adopt a more forceful carrot-and-stick approach to bringing about a change in the culture of the biomedical imaging community, as we are convinced that more open access will pay big dividends.

8 LOOKING BEYOND THE FIVE YEAR HORIZON – "SUPERCOMPUTING" AND MEDICAL IMAGING

Recognizing the gigantic size and intrinsic conservatism of the medical (and medical imaging) community, most of this report is rightly limited to 'the art of the possible' – recommendations for incremental change that leverage off of prior art, on a 5 year time scale. However, we would be remiss if we did not make at least some attempt at a more radical 'futurism', outlining what sorts of advances could, in principle, be achieved by major investments in paradigm-breaking technologies.

It is not by happenstance that practicing radiologists are fully trained as physicians before they acquire any specialized training in medical imaging and image interpretation. As a physician, the radiologist has peered-at, poked, palpated, prodded, pondered, and in many cases dissected the tissues and organs whose images will fill the rest of his or her career. The result of this early training is that the radiologist has a mental model not just in image space, but in the underlying 'real' space of anatomy and physiology.

This is a profound point: The radiologist is able, with an ease that comes from training and experience, to 'filter' the huge space of all possible (distorted, noisy, imperfect, ...) images into the large, but tractable, space of anatomically and physiologically possible situations (conditions, processes, syndromes, diseases, ...). This is a huge, and necessary, dimensional reduction in the image interpretation problem. While no two individuals, even normal individuals, are identical at the image level, the filter of understanding the 'laws' of physiology, etc., enables the radiologist to see non-identical images as belonging to common equivalence classes.

8.1 Using Models to Reduce the Dimensionality of the Image Analysis Problem

A national stretch goal for computation in support of medical imaging would be to develop a level of computer understanding, based on an underlying physically simulated model, comparable to that of an experienced radiologist.

This is not as crazy as it sounds. We will not be asking the computer to understand medicine, or to make diagnostic judgments, but only to understand anatomical, physical, mechanical, and possibly chemical properties of the tissues of the human body: mechanical and elastic properties, fluid flows (both free and diffusive flows), stress and strain relationships, and so on; and to be able to model these relationships in the presence of constraints imposed by the data of medical images. Although perhaps harder in practice, this is not different in principle from the problem of modeling the detonation behavior of a nuclear weapon, constrained by the image data of nuclear and non-nuclear tests – a problem in which the nation has invested several billions of dollars and with highly successful return.

According to the data provided by Mark Ellisman (8), a brain of 1500 cm^3 can yield an enormous amount of data. For micron-scale spatial resolution and 3 bytes/pixel, a single full-brain image would require 4.5 Petabytes of data storage.

If it is indeed possible to eventually image at this level, then there is clearly a data storage problem that cannot easily be managed. There is good reason to believe that disk drives will top out at a few TB each. Let's imagine that 10TB is a reasonable terminal disk size. Then a color 1 μm image would require 450 such disk drives. The time to read one of these disk drives at 1GB/s (which is roughly 20 times what can be done today) is 10^4 seconds, a little over three hours. If the data were striped across all disks, and assuming you could build a memory with that kind of bandwidth (current memories are a few GB/s at best, the ESS doing parallel memory

accesses is 32 GB/s), then it could be done in about three hours. This also represents the best case scenario for writing the data. If the data were placed sequentially on the disks, then it would take 56 days to read a single data set. This of course assumes sequential access to the data, which is the highest bandwidth form of access. Smaller accesses are possible, but the data needs to be structured in such a way that the volume to be extracted can be done easily and in parallel, it could easily degenerate to close to the worst case sequential scenario even in the case of highly striped disks (a small read from a single disk in each stripe).

It seems unlikely that 1 μm resolution is likely to occur in the near future. What we showed is that a single brain at 1 μm resolution was equal to the next generation ASCI computer in terms of disk storage, and would exceed the memory of that computer by orders of magnitude. It is interesting to note that 1 million brains at 1 mm resolution (current MRI resolution) is 4.5TB, which is manageable. It is important to note that this is to store an image at 1 mm resolution, not the data used to construct that image. The data used to construct at MRI image using a single coil is approximately 2GB, looking to the near future where arrays of 16 coils will be used the data grows to 32GB. If we return to the database of 1 million brains, then this means that from 2PB to 32PB of data must be stored for 1 mm resolution.

8.2 Tracking Changes in Each Patient

Since in the future a single individual will be imaged multiple times in a lifetime, the computer also needs to 'understand' (i.e., have available in a form able to be manipulated as a physical model) some areas of developmental biology. For example: The geometry of the brain's cortical folds are to some extent common to all normal individuals, and to some extent random (as the growing brain is packed, in mechanical equilibrium, into the growing skull). With a database of all previous images of an individual, *and* with a real-time mechanical model of how brain tissue responds to mechanical stresses, the computer will disambiguate normal small changes from (e.g.)

incipient tumor growth.

Particularly when the common clinical practice evolves to include periodic full body scans, it should be fairly straightforward to classify each individual's physiological anomalies so that the detection of anomalies or pathologies is not confounded by benign anatomical anomalies.

There already exist pilot efforts in disparate fields that are steps towards reaching this stretch goal. For example, the "Cardiome" project(10) is attempting to develop an integrated model of the heart, incorporating mechanical simulation, fluid flow, neuro-electrical behavior, and so forth. In the entirely different field of computer animation, there exist skeletal models of the human body, with mechanically realistic representations of muscle, draped skin, and so forth. These are computed according to the actual laws of physics so as to achieve realistic animations.

What we need is the 'full body' model – not at the molecular or biochemical level, but at the level of reproducing all the features that are accessible to medical imaging. Further, we need this model to be not just a 'forward' model (the kind that can predict appearance given state) but also to have the right computational 'hooks' in it to be usable as a 'backward' model, whereby state can be inferred by images. It would be an important part of the research agenda to define exactly what these hooks should be: This would be research combining computer science with medical expertise on the complete catalog of conditions that one expects to diagnose by imaging.

Baysian statistics is already an integral, if subconscious, part of medical image analysis and interpretation. Physicians assess the likelihood of different interpretations of an image based in large part on an appraisal of prior probabilities, drawn from case histories and other clinical information. This would have to be formalized and incorporated into the scheme described here, and this will require considerable development and testing.

8.3 Taking Steps in This Direction

This section has considered an approach in which the number of degrees of freedom in medical image analysis is radically reduced, by imposing physical and physiological constraints via a full computational model. This is essentially what physicians do on a daily basis, and it is in principle within our reach, given a deep enough understanding coupled with adequate computing power. Moving in this direction would require a clear long-term view on the part of the agencies, coupled with a staged program of research and development.

Existing computing resources within the DOE complex could be brought to bear on example problems of limited scope, and computational scaling performance could be explored and evaluated. In addition, a program of aggressive algorithmic and model development would be required.

It is important to not have the scope of our vision limited by our present computational capabilities. It seems to us inevitable that the capabilites that are presently available in state-of-the-art supercomputers will eventually migrate to the desktop. It's only a matter of *when* this will occur. It makes sense to be prepared to exploit the continued evolution of available computational power, and to remain open to revolutionary rather than evolutionary developments.

9 RECOMMENDATIONS AND CONCLUSIONS

We have summarized our view, on a 5 year time scale, of the computational requirements for medical imaging in the chart shown in Figure 4.

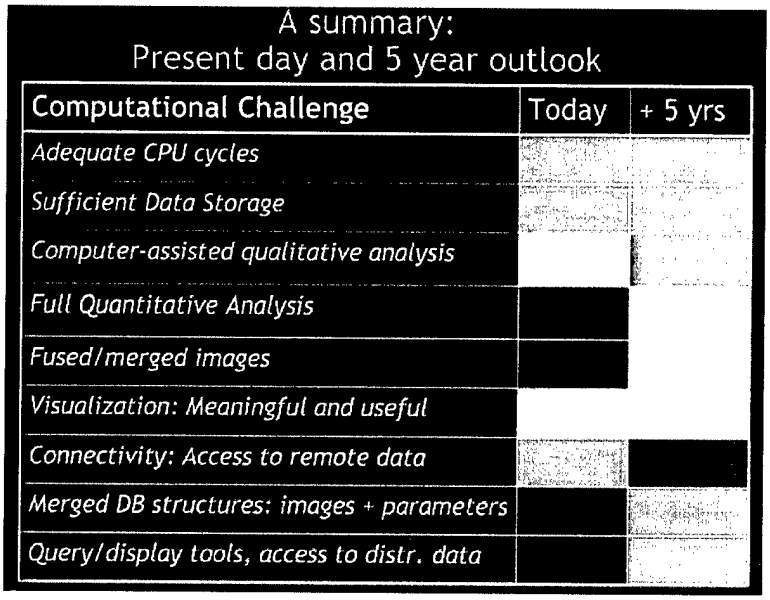

Figure 4: Summary of Computing in Support of Medical Imaging. The chart shows the JASON appraisal of the status of various computational needs for medical imaging. Green sections indicate items where needs are well met, yellow segments merit concern, and red segments are areas of serious deficiency.

Our recommendations were presented briefly in the Executive Summary, and are repeated here with somewhat more elaboration. We consider these recommendations to be high-leverage opportunities. Some will provide near-term dividends. Others represent our attempt to anticipate bottlenecks that are likely to arise further into the future.

1. Implement the BISTI report recommendations. In particular their recommendation number 4, pertaining to the availability of a hierarchy of computing platforms for the biological community, is essential to continued progress in biomedical imaging. The legendary benefits of

Moore's Law only accrue if new hardware is procured on a timely basis. An important aspect of this is to provide resources to supply the biomedical imaging community with a hierarchy of computing tools, ranging from desktop systems to supercomputer facilities. Equally important is providing funding to acquire mass storage capacity. These procurements, however, must go hand-in-hand with the definition and adoption of metadata standards, in order to ensure that the imaging data and derived products will be sustainable with the inevitable turnover in computer hardware, operating systems, and software.

2. Calibrate! The lack of credible geometrical registration hampers image fusion, and uncalibrated absorbtion or other information hampers the quantitative interpretation of biomedical images. We encourage working towards distribution of 3-d standards for geometrical registration frames, incorporating calibration as in integral part of each measurement, and appending the calibration information to all raw data files. In addition, the actual measured physical parameters (transmission, density...) should be measured, to the extent possible, in calibrated physical units. This also will support moving towards the incorporation of meaningful uncertainties as an integral part of biomedical imaging data.

3. Cultivate an open-access and open-source approach to biomedical imaging data sets and analysis algorithms. There are significant cultural impediments within the biomedical imaging community to the sharing of images and algorithms. The current NIH standards for data access stand in stark contrast to the common practice in other disciplines. This includes even the publication norms of other branches of the life sciences, such as genetic sequence data being made public is a condition of publication of research results. Furthermore, there are no common set of 'test problems' against which new algorithms can be tested. We advocate addressing these issues by nurturing the sharing of both code and data. One specific possibility is given in the following recommendation.

4. Establish an open ("BioLena") data set, which all researchers can use to test algorithms and techniques. We have in mind a set of images akin to those used by the computer imaging community, which are used as test images in essentially all research on algorithms and image processing. Implementing prototype metadata standards, NIBIB could act as curators, allowing apples-to-apples comparisons and industry standard test problems. We propose data sets (both raw and processed) that are drawn from each of the biomedical imaging modalities. We also advocate encouraging researchers to post, for open access, images that result from applying their new analysis or reduction algorithms. This will promote progress in metadata standards as well as providing a mechanism for quantitative, scientific, comparison of different algorithms.

5. Promote computer-assisted qualitative analysis of biomedical images in the clinical arena. This intermediate step strikes us an achievable near-term goal along the path towards eventual automated quantitative analysis of biomedical images. We think it is relatively straightforward to use existing technology to present the physician with not only the clinical images from a single patient, but also with a mosaic of images from comparable cases, along with their histories and outcomes. This may require some work to deal with patient confidentiality issues, but that strikes us as a tractable problem. One could also imagine an interactive image display system that is optimized to assist with differential diagnosis challenges.

6. Develop appropriate database technology, and select and evaluate demonstration projects. We see the database challenges associated with biomedical image exploitation as a major technical bottleneck in the coming years, but one which can be somewhat averted if appropriate steps are taken now. A particular topic for long term research is feature-based image queries, in which the step of parameterizing image features is not an explicit stage of image analysis, which produces an intermediate data catalog that is the basis for comparisons.

7. Establish a succession of "Grand Challenge Problems in Biomedical Imaging" to stimulate technical progress on the roadblock issues listed above. This approach has served the DOE community well in the past. A clear example is the success of the protein folding competitions which are now a staple of computational molecular biology. These challenges can also be crafted to galvanize collaborations between the biomedical community, mathematicians and computational and database scientists. Example problems include:

   - Map-the-Phantom – Construct a full-scale anatomical model (thoracic, cerebral?) and invite teams to acquire images and then provide their best quantitative, distortion-corrected reconstruction of the interior structure of the model. Kudos to those who produce the highest fidelity data set.
   - Multi-scale integration – Functional imaging of a biological process, from molecular to physiology. Examples are the cardio and brain efforts already under way. This will promote the eventual adoption of cellular and molecular imaging as clinical techniques.
   - Time-to-solution challenge – Pick an imaging methodology and problem. Points for whoever can port their analysis toolkit to a standard platform and get an acceptable answer the fastest. Also award points for the "best" answer.
   - Quantitative Change Detection Challenge – Given a temporal sequence of images, some with actual clinical data and others with features inserted "by hand", identify and quantify the evolution of the changes. We consider this as a tractable aspect of quantitative biomedical image analysis, rather than trying to solve the more general problem of recognizing and characterizing all features in an arbitrary image.
   - Multimode image integration challenge – Produce the best registered set of images, with common data structures and access tools, from images obtained with diverse methods. There is also an opportunity here to promote the *joint* analysis of raw data, rather

than just merging images after all processing has been done.

- Best merged Image-plus-catalog data structure, with query tools and comparison metrics for images. This will move the field in the direction of merged data entities, and will help lay important groundwork for development of metadata standards.

8. Begin the process of considering the potential of using what we presently consider super-computing in the biomedical imaging arena. Today's supercomputer is tomorrow's desktop machine, and this may open up totally new approaches to the interpretation of biomedical images.

10 ACKNOWLEDGMENTS

We are very grateful to the speakers listed in Table 1, who were very generous with their time, and very patient with our questions.

References

[1] *Biomedical Information Science and Technology Initiative Report*, Working Group on Biomedical Computing, `http://www.nih.gov/about/director/060399.htm` (1999)

[2] *Opportunities for Biomedical Research and the NIH through High Performance Computing and Data Management* Prepared by the Coalition for Advanced Scientific Computing, `http://www.casc.org`, (2003)

[3] *Transforming Health Care Through Information Technology*, President's Information Technology Advisory Committee, National Coordination Office for Information Technology Research and Development, (2001).

[4] *Sharing Images*, Vannier and Summers, editorial, Radiology, in press (2003).

[5] *NIH Data Access Policy*, `http://grants1.nih.gov/grants/guide/notice-files/NOT-OD-03-032.html`, (2003)

[6] L. Frank, JASON presentation, 2003

[7] C. Alard and R. Lupton, A Method for Optimal Image Subtraction, Astrophysical Journal **503** 325, 1998

[8] M. Ellisman, JASON Presentation, 2003.

[9] `http://adc.gsfc.nasa.gov/ gass/linsky/linsky.html`

[10] J B Bassingthwaighte, H Qian and Z Li, The Cardiome Project: An integrated view of cardiac metabolism and regional mechanical function, Adv. Exp. Med. Biol., **471**, 541-553, 1999

DISTRIBUTION LIST

Director of Space and SDI Programs
SAF/AQSC
1060 Air Force Pentagon
Washington, DC 20330-1060

CMDR & Program Executive Officer
U S Army/CSSD-ZA
Strategic Defense Command
PO Box 15280
Arlington, VA 22215-0150

DARPA Library
3701 North Fairfax Drive
Arlington, VA 22203-1714

Department of Homeland Security
Attn: Dr. Maureen McCarthy
Science and Technology Directorate
Washington, DC 20528

Assistant Secretary of the Navy
(Research, Development & Acquisition)
1000 Navy Pentagon
Washington, DC 20350-1000

Principal Deputy for Military Application [10]
Defense Programs, DP-12
National Nuclear Security Administration
U.S. Department of Energy
1000 Independence Avenue, SW
Washington, DC 20585

Superintendent
Code 1424
Attn: Documents Librarian
Naval Postgraduate School
Monterey, CA 93943

DTIC [2]
8725 John Jay Kingman Road
Suite 0944
Fort Belvoir, VA 22060-6218

Strategic Systems Program
Nebraska Avenue Complex
287 Somers Court
Suite 10041
Washington, DC 20393-5446

Headquarters Air Force XON
4A870 1480 Air Force Pentagon
Washington, DC 20330-1480

Defense Threat Reduction Agency
Attn: Dr. Arthur T. Hopkins [12]
8725 John J. Kingman Rd
Mail Stop 6201
Fort Belvoir, VA 22060-6201

IC JASON Program [2]
Chief Technical Officer, IC/ITIC
2P0104 NHB
Central Intelligence Agency
Washington, DC 20505-0001

JASON Library [5]
The MITRE Corporation
3550 General Atomics Court
Building 29
San Diego, California 92121-1122

U. S. Department of Energy
Chicago Operations Office Acquisition and
Assistance Group
9800 South Cass Avenue
Argonne, IL 60439

Dr. Jane Alexander
Homeland Security: Advanced Research
Projects Agency, Room 4318-23
7th & D Streets, SW
Washington, DC 20407

Dr. A. Michael Andrews
Director of Technology
SARD-TT - Room 3E480
Research Development Acquisition
103 Army Pentagon
Washington, DC 20310-0103

Dr. William O. Berry
Director, Basic Research ODUSD(ST/BR)
4015 Wilson Blvd
Suite 209
Arlington, VA 22203

Dr. Albert Brandenstein
Chief Scientist
Office of Nat'l Drug Control Policy Executive
Office of the President
Washington, DC 20500

Ambassador Linton F. Brooks
Under Secretary for Nuclear Security/
Administrator for Nuclear Security
1000 Independence Avenue, SW
NA-1, Room 7A-049
Washington, DC 20585

Dr. Darrell W. Collier
Chief Scientist
U. S. Army Space & Missile Defense Command
PO Box 15280
Arlington, VA 22215-0280

Dr. James F. Decker
Principal Deputy Director
Office of the Director, SC-1
Room 7B-084
U.S. Department of Energy
1000 Independence Avenue, SW
Washington, DC 20585

Dr. Patricia M. Dehmer [5]
Associate Director of Science for Basic Energy
Sciences, SC-10/Germantown Building
U.S. Department of Energy
1000 Independence Ave., SW
Washington, DC 20585-1290

Ms. Shirley A. Derflinger [15]
Technical Program Specialist
Office of Biological & Environmental Research
SC-70/Germantown Building
U.S. Department of Energy
1000 Independence Ave., SW
Washington, D.C. 20585-1290

Dr. Martin C. Faga
President and Chief Exec Officer
The MITRE Corporation
Mail Stop N640
7515 Colshire Drive
McLean, VA 22102

Mr. Dan Flynn [5]
Program Manager
DI/OTI/SAG
5S49 OHB
Washington, DC 20505

Ms. Nancy Forbes
Senior Analyst
DI/OTI/SAG 5S49 OHB
Washington, DC 20505

Dr. Paris Genalis
Deputy Director
OUSD(A&T)/S&TS/NW
The Pentagon, Room 3D1048
Washington, DC 20301

Mr. Bradley E. Gernand
Institute for Defense Analyses
Technical Information Services, Room 8701
4850 Mark Center Drive
Alexandria, VA 22311-1882

Dr. Lawrence K. Gershwin
NIC/NIO/S&T
2E42, OHB
Washington, DC 20505

Brigadier General Ronald Haeckel
U.S. Dept of Energy
National Nuclear Security Administration
1000 Independence Avenue, SW
NA-10 FORS Bldg
Washington, DC 20585

Dr. Theodore Hardebeck
STRATCOM/J5B
Offutt AFB, NE 68113

Dr. Robert G. Henderson
Director, JASON Program Office
The MITRE Corporation
7515 Colshire Drive
Mail Stop T130
McLean, VA 22102

Dr. Charles J. Holland
Deputy Under Secretary of Defense Science
& Technology
3040 Defense Pentagon
Washington, DC 20301-3040

Dr. Bobby R. Junker
Office of Naval Research
Code 31
800 North Quincy Street
Arlington, VA 22217-5660

Dr. Andrew F. Kirby
DO/IOC/FO
6Q32 NHB
Central Intelligence Agency
Washington, DC 20505-0001

Dr. Anne Matsuura
Army Research Office
4015 Wilson Blvd
Tower 3, Suite 216
Arlington, VA 22203-21939

Mr. Gordon Middleton
Deputy Director
National Security Space Architect
PO Box 222310
Chantilly, VA 20153-2310

Dr. Julian C. Nall
Institute for Defense Analyses
4850 Mark Center Drive
Alexandria, VA 22311-1882

Dr. C. Edward Oliver [5]
Associate Director of Science for Advanced
Scientific Computing Research
SC-30/Germantown Building
U.S. Department of Energy
1000 Independence Avenue, SW
Washington, DC 20585-1290

Mr. Raymond L. Orbach
Director, Office of Science
U.S. Department of Energy
1000 Independence Avenue, SW
Route Symbol: SC-1
Washington, DC 20585

Dr. Ari Patrinos [5]
Associate Director of Science for Biological
and Environmental Research
SC-70/Germantown Building
US Department of Energy
1000 Independence Avenue, SW
Washington, DC 20585-1290

Dr. John R. Phillips
Chief Scientist, DST/CS
2P0104 NHB
Central Intelligence Agency
Washington, DC 20505-0001

Records Resource
The MITRE Corporation
Mail Stop D460
202 Burlington Road, Rte 62
Bedford, MA 01730-1420

Dr. John Schuster
Submarine Warfare Division
Submarine, Security & Tech Head (N775)
2000 Navy Pentagon, Room 4D534
Washington, DC 20350-2000

Dr. Ronald M. Sega
DDR&E
3030 Defense Pentagon , Room 3E101
Washington, DC 20301-3030

Dr. Alan R. Shaffer
Office of the Defense Research and Engineering
Director, Plans and Program
3040 Defense Pentagon, Room 3D108
Washington, DC 20301-3040

Mr. Frank Spagnolo
Advanced Systems & Technology
National Reconnaissance Office
14675 Lee Road
Chantilly, Virginia 20151

Mr. Anthony J. Tether
DIRO/DARPA
3701 N. Fairfax Drive
Arlington, VA 22203-1714

Dr. Bruce J. West
FAPS - Senior Research Scientist
Army Research Office
P. O. Box 12211
Research Triangle Park, NC 27709-2211

Dr. Linda Zall
Central Intelligence Agency
DS&T/OTS
3Q14, NHB
Washington, DC 20505-00